Computers and Mathematics with Applications 61 (2011) 1407–1418

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

A modified smoothing and regularized Newton method for monotone second-order cone complementarity problems☆

Linjie Chen, Changfeng Ma*

School of Mathematics and Computer Sciences, Fujian Normal University Fuzhou, 350007, China

ARTICLE INFO

Article history:
Received 13 May 2010
Received in revised form 7 January 2011
Accepted 9 January 2011

Keywords:
Second-order cone complementarity problem
Smoothing method
Regularization
Convergence analysis
Numerical results

ABSTRACT

In this paper, we propose a globally and quadratically convergent Newton-type algorithm for solving monotone second-order cone complementarity problems (denoted by SOCCPs). This algorithm is based on smoothing and regularization techniques by incorporating smoothing Newton's method. Many Newton-type methods with smoothing and regularization techniques have been studied for solving nonlinear complementarity problems (NCPs) and box constrained variational inequalities (BVIs). Our algorithm is regarded as an extension of those methods to SOCCP. However, it is different from the existing methods, because we solve SOCCP by treating both the smoothing parameter μ and the regularization parameter ε as independent variables. In addition, numerical experiments indicate that the proposed method is quite effective.

1. Introduction

The second-order cone (SOC) $\mathcal{K}^n$ in R^n $(n \geq 1)$ is defined to be

$$\mathcal{K}^n = \{(x_1, x_2^T)^T \in R \times R^{n-1} \,|\, \|x_2\| \leq x_1\},$$

where $\|\cdot\|$ denotes the Euclidean norm. It is also called the Lorentz cone, because of the special ice-cream shape when $n = 3$. Recently, the second-order cone complementarity problem (SOCCP) has been the focus of several studies [1–6]. SOCCP is a wide class of problems containing the nonlinear complementarity problem (NCP) and the second-order cone programming problem (SOCP) [2].

The second-order cone complementarity problem (SOCCP) [1] is to find $(x, y, \zeta) \in R^n \times R^n \times R^l$ such that

$$x \in \mathcal{K},\ y \in \mathcal{K}, x^T y = 0, \quad F(x, y, \zeta) = 0, \tag{1.1}$$

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m} \subset R^n, \tag{1.2}$$

where $F : R^n \times R^n \times R^l \to R^n \times R^l$ is a continuously differentiable mapping and $\mathcal{K} \subset R^n$, with $l \geq 0, m, n_1, n_2, \ldots, n_m \geq 1$ and $n_1 + n_2 + \cdots + n_m = n$. If $n_1 = n_2 = \cdots = n_m = 1$ and $F(x, y, \zeta) = f(x) - y$ then the SOCCP becomes NCP. In this paper, we focus on the special SOCCP: find $(x, y) \in R^n \times R^n$ such that

$$x \in \mathcal{K}, \quad y = F(x) \in \mathcal{K}, \quad \langle x, y\rangle = 0, \tag{1.3}$$

where $\langle\cdot, \cdot\rangle$ denotes the Euclidean inner product, and F is a continuously differentiable mapping from R^n to R^n. Note that the complementarity condition on $\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m}$ can be decomposed into complementarity conditions on each

☆ This project is supported by the National Natural Science Foundation of China (Grant No. 11071041) and the Fujian National Natural Science Foundation (No. 2009J01002).
* Corresponding author. Tel.: +86 13763827962.
E-mail address: macf@fjnu.edu.cn (C. Ma).

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.01.009

$\mathcal{K}^{n_i}$ $(i = 1, 2, \ldots, m)$, that is

$$x \in \mathcal{K},\ y \in \mathcal{K}, \quad x^T y = 0, \Longleftrightarrow$$
$$x^i \in \mathcal{K}^i,\ y^i \in \mathcal{K}^i, \quad (x^i)^T y^i = 0,\ (i = 1, 2, \ldots, m),$$

where $x = (x^1, x^2, \ldots, x^m) \in R^{n_1} \times R^{n_2} \times \cdots \times R^{n_m}$ and $y = (y^1, y^2, \ldots, y^m) \in R^{n_1} \times R^{n_2} \times \cdots \times R^{n_m}$.

It is known that SOCCP can be reduced to an SDCP [4] that is, for any $x = (x_1, x_2) \in R \times R^{n-1}$, we have $x \in \mathcal{K}^n$ if and only if

$$L_x := \begin{pmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{pmatrix},$$

where I denotes the identity matrix, and the matrix L_x is positive semidefinite [7]. Moreover, L_x is positive definite (and hence invertible) if and only if $x \in \mathrm{int}(\mathcal{K}^n)$. However, this reduction increases the problem dimension from n to $n(n+1)/2$.

In the rest of this section, we review the Jordan product and the spectral factorization associated with a second-order cone, which plays a key role in analyzing the properties of merit functions for the SOCCP. Moreover, we introduce a kind of function associated with SOC based on the spectral factorization.

For any $x = (x_1, x_2),\ y = (y_1, y_2) \in R \times R^{n-1}$, we define their Jordan product [8] associated with $\mathcal{K}^n$ as

$$x \circ y := (\langle x, y\rangle, y_1 x_2 + x_1 y_2).$$

The identity element under this product is noted $e := (1, 0, \ldots, 0)^T \in R^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in R^n$. Hence, it is clear that there exists a unique vector denoted by $|x|$, so we have $x^2 = |x|^2$. Moreover, if $x \in \mathcal{K}^n$ then there exists a unique vector in $\mathcal{K}^n$, denoted by $x^{\frac{1}{2}}$, such that $(x^{\frac{1}{2}})^2 = x^{\frac{1}{2}} \circ x^{\frac{1}{2}} = x$. For each $x = (x_1, x_2) \in R \times R^{n-1}$, the determinant and the trace of x are defined by

$$\det(x) = x_1^2 - \|x_2\|^2 \quad \text{and} \quad \mathrm{tr}(x) = 2x_1.$$

We next review the spectral factorization of vectors in R^n associated with $\mathcal{K}^n$ [1]. For any vector $x = (x_1, x_2) \in R^n$ admits a spectral factorization, associated with $\mathcal{K}^n$, of the form

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \tag{1.4}$$

where λ_1 and λ_2 are the spectral values given by

$$\lambda_i = x_1 + (-1)^i \|x_2\|, \quad i = 1, 2, \tag{1.5}$$

and $u^{(1)}$ and $u^{(2)}$ are the spectral vectors given by

$$u^{(i)} = \begin{cases} \dfrac{1}{2}\left(1, (-1)^i \dfrac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \dfrac{1}{2}\left(1, (-1)^i w_2\right), & \text{if } x_2 = 0, \end{cases} \tag{1.6}$$

for $i = 1, 2$, with w_2 being any vector in R^{n-1} satisfying $\|w_2\| = 1$. If $x_2 \neq 0$, the decomposition (1.4) is unique.

According to the spectral factorization associated with a second-order cone, we introduce a kind of function associated with SOC.

For any differentiable convex function $\hat{g} : R \to R$ satisfying

$$\lim_{\alpha \to -\infty} \hat{g}(\alpha) = 0, \qquad \lim_{\alpha \to \infty} (\hat{g}(\alpha) - \alpha) = 0, \quad 0 < \hat{g}'(\alpha) < 1,$$

for example, $\hat{g}_1(\alpha) = (\sqrt{\alpha^2 + 4} + \alpha)/2$ and $\hat{g}_2(\alpha) = \ln(e^\alpha + 1)$ satisfy the conditions. Furthermore, using $\hat{g}(\cdot)$ we define a function on R^n associated with $\mathcal{K}^n$ $(n \geq 1)$ by

$$g(x) = \hat{g}(\lambda_1) u^{(1)} + \hat{g}(\lambda_2) u^{(2)} \quad \text{for any } x = (x_1, x_2) \in R \times R^{n-1}, \tag{1.7}$$

where λ_1, λ_2 and $u^{(1)}, u^{(2)}$ are the spectral values and vectors of x (see (1.4)–(1.6)). Definition (1.7) is unambiguous when $x_2 \neq 0$ since $u^{(1)}, u^{(2)}$ are unique, when $x_2 = 0$, we see from (1.5) and (1.6) that $g(x) = \hat{g}(x_1)e$ so the definition also unambiguous again. Specially, the cases of $g(x) = x^{\frac{1}{2}}$, is discussed in the book of Faraut and Korányi [8].

This paper is organized as follows. In Section 2, first we show the well-known merit function of SOCCP that plays a key role in the subsequent analysis. Then we introduce smoothing and regularization methods, and consider the boundedness of level sets that plays an important role in the global convergence of a descent method. In Section 3, we give an algorithm for solving the SOCCP. Convergence analysis of the algorithm is given in Section 4. Some numerical results are reported in Section 5.

Throughout this paper, we use the following notation: R^n $(n \geq 1)$ denotes the space of n-dimensional real column vectors and ^{T} denotes transpose. We write $x = (x_1, x_2)$ for $(x_1, x_2^T)^T$, R_+ and R_{++} denote the nonnegative and positive reals, respectively, therefore $\mathcal{K}^1$ denotes R_+. For any $x, y \in R^n$, the Euclidean inner product denotes $\langle x, y\rangle = x^T y$, $\|\cdot\|$ denotes the 2-norm defined by $\|x\| := \sqrt{x^T x}$.

2. Smoothing and regularization

In the context of SOCCP, for simplicity, we assume $\mathcal{K} = \mathcal{K}^n$. In order to solve the SOCCP, we introduce a convenient merit function $\Psi : R^n \times R^n \to R_+$, and $\Psi(x, y) = 0$ if and only if (x, y) is a solution of SOCCP (1.3). To this aim we firstly construct a function $\psi : R^n \times R^n \to R^n$ satisfying

$$\psi(x, y) = 0 \Leftrightarrow \langle x, y \rangle = 0, \quad x \in \mathcal{K},\ y \in \mathcal{K}. \tag{2.1}$$

In this paper, we regard the well-known Fischer–Burmeister function $\phi_{\rm FB} : R^n \times R^n \to R^n$ as function ψ, which defined by

$$\phi_{\rm FB}(x, y) = x + y - \sqrt{x^2 + y^2}. \tag{2.2}$$

By using such a function, we define $H_{\rm FB} : R^n \times R^n \to R^{2n}$ by

$$H_{\rm FB}(x, y) = \begin{pmatrix} \phi_{\rm FB}(x, y) \\ F(x) - y \end{pmatrix}.$$

It shows that SOCCP (1.3) is equivalent to the equation $H_{\rm FB}(x, y) = 0$. Furthermore, we can define function $\Psi_{\rm FB}$ by

$$\Psi_{\rm FB}(x, y) := \frac{1}{2}\|H_{\rm FB}(x, y)\|^2 = \frac{1}{2}\|\phi_{\rm FB}(x, y)\|^2 + \frac{1}{2}\|F(x) - y\|^2. \tag{2.3}$$

Therefore, this function $\Psi_{\rm FB}$ defined by (2.3) can serve as a merit function for SOCCP (1.3).

From definition (2.2) of $\phi_{\rm FB}$, we consider a type of smoothing function $\phi(\mu, x, y) : R \times R^n \times R^n \to R^n$. Fukushima et al. [1] extended the Fischer–Burmeister function from NCP to SOCCP. That is,

$$\phi(\mu, x, y) = x + y - (x^2 + y^2 + 2\mu^2 e)^{\frac{1}{2}}, \tag{2.4}$$

which function is viewed as a smoothing approximation function of $\phi_{\rm FB}$.

Therefore, function $\tilde{\Psi} : R_{++} \times R^n \times R^n \to R_+$ given by

$$\tilde{\Psi}(\mu, x, y) := \frac{1}{2}\|\phi(\mu, x, y)\|^2 + \frac{1}{2}\|F(x) - y\|^2,$$

serves as a smoothing approximation function of the merit function $\Psi_{\rm FB}$. As we know, the boundedness of the level sets $\mathcal{L}_\sigma := \{(x, y) | \tilde{\Psi}(\mu, x, y) \le \sigma\}$ for all $\sigma \in R_+$ can guarantee that a sequence generated by an appropriate descent method has an accumulation point. And the level-boundedness of the objective function plays an important role in the convergence of a descent method [2]. If the function involved in the SOCCP is strongly monotone, then the merit function $\tilde{\Psi}$ is level-bounded. However, strong monotonicity is quite a severe condition. To be amenable to a monotone problem, we employ a regularization method.

Let $F(x)$ be a monotone function, that is, for any $(x, y) \in R^n \times R^n$, $(x - y)^T[F(x) - F(y)] \ge 0$ always hold. We can define a function $F_\varepsilon := F(x) + \varepsilon x$ with a parameter $\varepsilon > 0$. Furthermore, a solution of the original SOCCP is obtained by taking the limit $\varepsilon \downarrow 0$. Therefore,

$$\begin{aligned} (x - y)^T\big(F_\varepsilon(x) - F_\varepsilon(y)\big) &= (x - y)^T\big(F(x) - F(y) + \varepsilon(x - y)\big) \\ &= \varepsilon\|x - y\|^2 + (x - y)^T\big(F(x) - F(y)\big) \\ &\ge \varepsilon\|x - y\|^2. \end{aligned}$$

According to the definition of a strong monotone function, we say that F_ε is strongly monotone for any $\varepsilon > 0$.

For any $z = (\mu, \varepsilon, x, y) \in R_{++} \times R_+ \times R^n \times R^n$, define function H and Ψ by

$$H(z) := H(\mu, \varepsilon, x, y) = \begin{pmatrix} \mu \\ \varepsilon \\ \phi(\mu, x, y) \\ F(x) + \varepsilon x - y \end{pmatrix} \tag{2.5}$$

and

$$\begin{aligned} \Psi(z) &:= \Psi(\mu, \varepsilon, x, y) = \frac{1}{2}\|H(\mu, \varepsilon, x, y)\|^2 \\ &= \frac{1}{2}(\mu^2 + \varepsilon^2) + \frac{1}{2}\|\phi(\mu, x, y)\|^2 + \frac{1}{2}\|F(x) + \varepsilon x - y\|^2, \end{aligned} \tag{2.6}$$

where $F(x)$ is a monotone function.

Because of the strong monotonicity of F_ε, the function $\Psi(\mu, \varepsilon, x, y)$ is level-bounded. In order to study the properties of function H, we first need the following couple technical lemmas from [1].

Lemma 2.1. *For any x, y in R^n and any $v \succ 0$, we have*

$$v^2 \succ x^2 + y^2 \Rightarrow (L_v - L_x)(L_v - L_y) \succ 0, L_v - L_x \succ 0,\ L_v - L_y \succ 0,$$

where $v \succ u$ denotes $v - u \in \operatorname{int}(\mathcal{K})$ and hence, $v \succ 0$ denote $v \in \operatorname{int}(\mathcal{K})$. Moreover, whenever "$\succ$" be replaced by "$\succeq$", this relation remains true, where $v \succeq u$ denotes $v - u \in \mathcal{K}$ and hence, $v \succeq 0$ denote $v \in \mathcal{K}$.

Lemma 2.2. *For any $\hat{g} : R \to R$ that is Fréchet-differentiable (respectively, continuously differentiable), the function $g : R^n \to R^n$ defined by (1.7) is Fréchet-differentiable (respectively, continuously differentiable) and its Jacobian at $z = (z_1, z_2) \in R \times R^{n-1}$ is given as*

$$\nabla g(z) = \begin{cases} g'(z_1)I & \text{if } z_2 = 0, \\ \begin{pmatrix} b, & cz_2^T/\|z_2\| \\ cz_2/\|z_2\|, & aI + (b-a)z_2z_2^T/\|z_2\|^2 \end{pmatrix} & \text{if } z_2 \neq 0, \end{cases} \tag{2.7}$$

where

$$a = \frac{\hat{g}(\lambda_2) - \hat{g}(\lambda_1)}{\lambda_2 - \lambda_1}, \qquad b = \frac{1}{2}\big(\hat{g}'(\lambda_2) + \hat{g}'(\lambda_1)\big), \qquad c = \frac{1}{2}\big(\hat{g}'(\lambda_2) - \hat{g}'(\lambda_1)\big)$$

with $\lambda_i = z_1 + (-1)^i\|z_2\|,\ i = 1, 2$. If $\hat{g}'(\alpha) > 0$ for all $\alpha \in R$, then $\nabla g(z)$ is positive definite for all $z \in R^n$. For example, we define $\hat{g}(\alpha) = \alpha^{\frac{1}{2}}$.

In the smoothing approach to solving (1.3), we solve a sequence of approximations to the original problem via an appropriate descent algorithm [9]. For the descent direction to be well defined and unique, the Jacobian matrix of $H(z)$ be invertible is so essential. As in the previous sections, we focus our analysis on the case of $\mathcal{K} = \mathcal{K}^n$ for simplicity.

Proposition 2.1. *Let $z = (\mu, \varepsilon, x, y) \in R_{++} \times R_+ \times R^n \times R^n$. From the definition of $H(z)$ and F_ε we can deduce the computing formulation of $\nabla H(z)$ as*

$$\nabla H(z)^T := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \phi'_\mu & 0 & \nabla_x\phi & \nabla_y\phi \\ 0 & x & F'(x) + \varepsilon I & -I \end{pmatrix}, \tag{2.8}$$

where ϕ is the smoothed Fischer–Burmeister function $\phi : R_{++} \times R^n \times R^n \to R^n$ given by (2.4) and it is continuously differentiable and the gradient of it with respect to μ, x, y, respectively, can be written as

$$\nabla_\mu\phi(\mu, x, y) = \phi'_\mu = -2\mu e L_v^{-1}, \tag{2.9}$$

$$\nabla_x\phi(\mu, x, y) = I - 2L_x\nabla g(z) = I - L_xL_v^{-1}, \tag{2.10}$$

$$\nabla_y\phi(\mu, x, y) = I - 2L_y\nabla g(z) = I - L_yL_v^{-1}, \tag{2.11}$$

where $u = x^2 + y^2 + 2\mu^2 e,\ v = u^{\frac{1}{2}}$, and $\nabla g(z)$ has the formula in Lemma 2.2 with $\hat{g}(\alpha) = \alpha^{\frac{1}{2}}$ for all $\alpha \in R_{++}$.

Proof. From (2.5) in a straightforward manner we can obtain

$$\nabla H(z)^T := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \phi'_\mu & 0 & \nabla_x\phi & \nabla_y\phi \\ 0 & x & F'(x) + \varepsilon I & -I \end{pmatrix}.$$

Notice that for any $\mu \in R_{++}$, the mapping $U(\mu, x, y) = x^2 + y^2 + 2\mu^2 e \succ 0$. Direct calculation yields

$$\begin{aligned} U(\mu, x, y) &= x \circ x + y \circ y + 2\mu^2 e \\ &= (\|x\|^2 + \|y\|^2 + 2\mu^2, 2(x_1x_2 + y_1y_2)). \end{aligned}$$

From the above equality we see that U is continuously differentiable and that

$$\begin{aligned} &\nabla_\mu U(\mu, x, y) = 4\mu e, \\ &\nabla_x U(\mu, x, y) = 2\begin{pmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{pmatrix} = 2L_x, \\ &\nabla_y U(\mu, x, y) = 2\begin{pmatrix} y_1 & y_2^T \\ y_2 & y_1 I \end{pmatrix} = 2L_y. \end{aligned} \tag{2.12}$$

From Lemma 2.2 we know that function $\hat{g}$ is continuously differentiable on R_{++}. Thus, $\psi(\mu, x, y) = \sqrt{x^2 + y^2 + 2\mu^2 e} = g \circ U$ is continuously differentiable and

$$\begin{aligned}\nabla_\mu \psi(\mu, x, y) &= \nabla_\mu U(\mu, x, y)\nabla g(U(\mu, x, y)),\\ \nabla_x \psi(\mu, x, y) &= \nabla_x U(\mu, x, y)\nabla g(U(\mu, x, y)),\\ \nabla_y \psi(\mu, x, y) &= \nabla_y U(\mu, x, y)\nabla g(U(\mu, x, y)).\end{aligned} \tag{2.13}$$

Now consider the term of $\nabla g(U(\mu, x, y))$.

As we know that $u \succ_{\mathcal{K}^n} 0$, there exists a scalar $\delta > 0$ such that for all $d \in R^n$ with $\|d\| < \delta$ yields $u + d \succ_{\mathcal{K}^n} 0$. For any such d, we let $p = (u + d)^{\frac{1}{2}} - v$ where $v = u^{\frac{1}{2}}$, so

$$d = (p + v)^2 - v^2 = p^2 + 2pv = p^2 + 2L_v p.$$

Since $v \succ_{\mathcal{K}^n} 0$, then L_v is invertible, let us multiply both sides of this equality by L_v^{-1} yields

$$L_v^{-1} d = 2p + L_v^{-1} p^2.$$

If $\|d\|$ is sufficiently small, then $\|p\| = O(\|d\|)$. Moreover,

$$g(u + d) - g(u) = p = \frac{1}{2} L_v^{-1} d - \frac{1}{2} L_v^{-1} p^2.$$

By the definition of Fréchet-differentiable, we have

$$\begin{aligned}\lim_{\|d\|\to 0} \frac{\|g(u + d) - g(u) - \frac{1}{2} L_v^{-1} d\|}{\|d\|} &= \lim_{\|d\|\to 0} \frac{\frac{1}{2}\|L_v^{-1} d^2\|}{\|d\|}\\ &\le \lim_{\|d\|\to 0} \frac{1}{2} \frac{\|L_v^{-1}\|\,\|d\|^2}{\|d\|} = \lim_{\|d\|\to 0} \frac{1}{2}\|L_v^{-1}\|\,\|d\| = 0,\end{aligned}$$

that is,

$$\nabla g(u) = \frac{1}{2} L_v^{-1}. \tag{2.14}$$

Using (2.12)–(2.14), yields

$$\nabla_\mu \phi = 2\mu e L_v^{-1}, \qquad \nabla_x \phi = I - L_x L_v^{-1}, \qquad \nabla_y \phi = I - L_x L_v^{-1}.$$

So we gain the required results. □

Proposition 2.2. *Suppose that F is a continuously differentiable monotone function $\phi : R \times R^n \times R^n \to R^n$ defined by (2.4). Then, for each $z := (\mu, \varepsilon, x, y) \in R_{++} \times R_+ \times R^n \times R^n$, the matrix $\nabla H(z)^T$ defined by (2.8) is invertible.*

Proof. Using Lemmas 2.1, 2.2 and Proposition 2.1 with $\hat{g}(\alpha) = \alpha^{\frac{1}{2}}$, we let $(\alpha, \beta, \Delta x, \Delta y) \in R \times R \times R^n \times R^n$ be a vector in the null space of $\nabla H(z)^T$. Then $\nabla H(z)^T(\alpha, \beta, \Delta x, \Delta y) = 0$, we will show that $\alpha = \beta = 0$, $\Delta x = \Delta y = 0$.

By Proposition 2.1, we have $\phi'_\mu(\mu, x, y) = 2\mu e L_v^{-1}$, $\nabla_x \phi(\mu, x, y)^T = I - L_v^{-1} L_x$ and $\nabla_y \phi(\mu, x, y)^T = I - L_v^{-1} L_y$, where $v = u^{\frac{1}{2}}$ and $u = x^2 + y^2 + 2\mu^2 e$. Then from the representation of $\nabla H(z)^T$, we have

$$\alpha = \beta = 0, \tag{2.15a}$$

$$\alpha \phi'_\mu + \nabla_x \phi(\mu, x, y)^T \Delta x + \nabla_y \phi(\mu, x, y)^T \Delta y = 0, \tag{2.15b}$$

$$\beta x + [F'(x) + \varepsilon I]\Delta x - \Delta y = 0. \tag{2.15c}$$

(2.15b) is equivalent to

$$(I - L_v^{-1} L_x)\Delta x + (I - L_v^{-1} L_y)\Delta y = 0; \tag{2.16}$$

we multiply both sides of this equality by L_v yields

$$(L_v - L_x)\Delta x + (L_v - L_y)\Delta y = 0.$$

From Lemma 2.1, we have $(L_v - L_x)(L_v - L_y) \succ 0$. Then, applying both sides on the left by $\Delta x^T (L_v - L_y)^{-1}$ yields

$$\Delta x^T (L_v - L_y)^{-1}(L_v - L_x)\Delta x + \Delta x^T \Delta y = 0. \tag{2.17}$$

Also using Lemma 2.1, we have $(L_v - L_x)(L_v - L_y) \succ 0$. Then we let $\xi = (L_v - L_y)^{-1}\Delta x$ hence the first term of (2.17) can be rewritten as

$$\Delta x^T (L_v - L_y)^{-1}(L_v - L_x)\Delta x = \xi^T (L_v - L_x)(L_v - L_y)\xi \ge 0.$$

Moreover, (2.15c) is equivalent to $\Delta y = (F'(x) + \varepsilon I)\Delta x$, so $\Delta x^T \Delta y = \Delta x^T(F'(x) + \varepsilon I)\Delta x \geq 0$. Since F is differentiable, F being monotone is equivalent to $F'(x) \succeq 0$ for all $x \in R^n$ [4] and $\varepsilon I \succ 0$ for all $\varepsilon \in R_{++}$. Thus, (2.17) implies that $\xi^T(L_v - L_x)(L_v - L_y)\xi = 0$, so we know that $\xi = 0$, then from the definition of ξ we have $\Delta x = 0$. Also since $I - L_v^{-1}L_y$ is invertible, (2.16) implies that $\Delta y = 0$.

So we say $(\alpha, \beta, \Delta x, \Delta y) = 0$. Thus, the null space of $\nabla H(z)^T$ comprises only the origin, so $\nabla H(z)^T$ is invertible. □

3. Algorithm

In the previous section, we have shown that if F is monotone, then for any $\mu > 0$ and $\varepsilon > 0$, the function $\Psi(z)$ with $z = (\mu, \varepsilon, x, y)$ defined by (2.6) is differentiable and level-bounded. In this section, we propose a modified algorithm based on smoothing and regularization techniques by incorporating Newton's method.

Choosing $\bar{\mu} \in R_{++}$, $\bar{\varepsilon} \in R_{++}$ and $\gamma \in (0, 1)$, such that $\gamma\bar{\mu} < 0.5$, $\gamma\bar{\varepsilon} < 0.5$. Let $\bar{z} = (\bar{\mu}, \bar{\varepsilon}, 0, 0)$ and $\Psi(z)$ be defined by (2.6). Moreover, we define $\rho : R_{++} \times R_{++} \times R^n \times R^n \to R_+$,

$$\rho(z) = \gamma \min\{1, \Psi(z)\}.$$

We also define a neighborhood

$$\Omega := \{z = (\mu, \varepsilon, x, y) \in R_{++} \times R_{++} \times R^n \times R^n \mid \mu \geq \rho(z)\bar{\mu}, \varepsilon \geq \rho(z)\bar{\varepsilon}\}.$$

Since $\rho(z) \leq \gamma < 1$, $(\bar{\mu}, \bar{\varepsilon}, x, y) \in \Omega$ for any $x, y \in R^n$.

Lemma 3.1. *$H(z) = 0 \Longleftrightarrow \rho(z) = 0 \Longleftrightarrow H(z) = \rho(z)\bar{z}$.*

Proof. By using the definition of $H(\cdot)$ and $\rho(\cdot)$ we have

$$H(z) = 0 \Leftrightarrow \rho(z) = 0 \quad \text{and} \quad \rho(z) = 0 \Rightarrow H(z) = \rho(z)\bar{z}.$$

Then we only need to prove that $H(z) = \rho(z)\bar{z} \Rightarrow \rho(z) = 0$.

Since $H(z) = \rho(z)\bar{z}$ so that $\mu = \rho(z)\bar{\mu}$, $\varepsilon = \rho(z)\bar{\varepsilon}$, $\phi(\mu, x, y) = 0$, $F(x) + \varepsilon x - y = 0$. From the definition of $\Psi(\cdot)$, $\rho(\cdot)$ and $\gamma\bar{\mu} < 0.5$, $\gamma\bar{\varepsilon} < 0.5$, hence

$$\begin{aligned}
\Psi(z) &= \frac{1}{2}\|H(z)\|^2 \\
&= \frac{1}{2}(\mu^2 + \varepsilon^2) + \frac{1}{2}\|\phi(\mu, x, y)\|^2 + \frac{1}{2}\|F(x) + \varepsilon x - y\|^2 \\
&= \frac{1}{2}(\mu^2 + \varepsilon^2) = \frac{1}{2}\rho(z)^2(\bar{\mu}^2 + \bar{\varepsilon}^2) \\
&\leq \frac{1}{2}\gamma^2(\bar{\mu}^2 + \bar{\varepsilon}^2) < 1.
\end{aligned}$$

Therefore,

$$\rho(z) = \gamma\Psi(z) = \frac{1}{2}\gamma\rho(z)^2(\bar{\mu}^2 + \bar{\varepsilon}^2). \tag{3.1}$$

If $\rho(z) \neq 0$, by (3.1) we have $\frac{1}{2}\gamma\rho(z)(\bar{\mu}^2 + \bar{\varepsilon}^2) = 1$, which, together with $\rho(z) \leq \gamma$, implies that

$$1 = \frac{1}{2}\gamma\rho(z)(\bar{\mu}^2 + \bar{\varepsilon}^2) \leq \frac{1}{2}\gamma^2(\bar{\mu}^2 + \bar{\varepsilon}^2),$$

which is in contradiction with $\gamma\bar{\mu} < 0.5$, $\gamma\bar{\varepsilon} < 0.5$. This contradiction completes our proof. □

Following Lemma 3.1, we will state the following algorithm.

Algorithm 3.1 (*A Smoothing–Regularization Newton Method*). *Step* 0: Choose parameters $\delta \in (0, 1)$, and $\sigma \in (0, 0.5)$. Choose $\bar{\mu} \in R_{++}$, $\bar{\varepsilon} \in R_{++}$ and $\gamma \in (0, 1)$, such that $\gamma\bar{\mu} < 0.5$, $\gamma\bar{\varepsilon} < 0.5$. Let $z_0 = (\bar{\mu}, \bar{\varepsilon}, x_0, y_0)$, where x_0, y_0 are arbitrary vectors in R^n. Set $k := 0$.

Step 1: If $\|H(z_k)\| = 0$ then stop. Otherwise, set $\rho_k = \rho(z_k)$.

Step 2: Solving

$$H(z_k) + \nabla H(z_k)^T \Delta z_k = \rho_k\bar{z}, \tag{3.2}$$

we obtain $\Delta z_k = (\Delta\mu_k, \Delta\varepsilon_k, \Delta x_k, \Delta y_k)$.

Step 3: Find the smallest nonnegative integer l_k, which satisfies the following inequality

$$\Psi(z_k + \delta^l \Delta z_k) \leq [1 - 2\sigma(1 - \gamma(\bar{\mu} + \bar{\varepsilon}))\delta^l]\Psi(z_k).$$

Let $z_{k+1} = z_k + \delta^{l_k}\Delta z_k$.

Step 4: Set $k := k + 1$. Go back to Step 1.

Proposition 3.1. *Algorithm 3.1 is well defined at the kth iteration. Furthermore, for some $\tilde{z} := (\tilde{\mu}, \tilde{\varepsilon}, \tilde{x}, \tilde{y}) \in R_{++} \times R_{+} \times R^n \times R^n$, there exist a closed neighborhood $\mathcal{N}(\tilde{z})$ of $\tilde{z}$ and a positive number $\tilde{\alpha} \in (0, 1]$ such that for any $z = (\mu, \varepsilon, x, y) \in \mathcal{N}(\tilde{z})$ and all $\alpha \in [0, \tilde{\alpha}]$, we have $\mu \in R_{++}$, $\varepsilon \in R_{+}$ which satisfy*

$$\Psi(z + \alpha \vartriangle z) \le [1 - 2\sigma(1 - \gamma(\bar{\mu} + \bar{\varepsilon}))\alpha]\Psi(z). \tag{3.3}$$

Proof. Since by Proposition 2.2, $\nabla H(\tilde{z})$ is invertible and $\tilde{\mu} \in R_{++}$, $\tilde{\varepsilon} \in R_{+}$, there exists a closed neighborhood $\mathcal{N}(\tilde{z})$ of $\tilde{z}$ such that for any $z = (\mu, \varepsilon, x, y) \in \mathcal{N}(\tilde{z})$ we have $z \in R_{++} \times R_{+} \times R^n \times R^n$ and $\nabla H(z)$ is invertible.

For any $z \in \mathcal{N}(\tilde{z})$, let $\Delta z = (\Delta\mu, \Delta\varepsilon, \Delta x, \Delta y) \in R \times R \times R^n \times R^n$ be the unique solution of the following equation

$$H(z) + \nabla H(z)^T \Delta z = \rho(z)\bar{z}. \tag{3.4}$$

And for any $\alpha \in [0, 1]$, notice that $p = (\mu, x, y)$, $\Delta p = (\Delta\mu, \Delta x, \Delta y)$, then define

$$g_p(\alpha) = \phi(p + \alpha\Delta p) - \phi(p) - \alpha\nabla\phi(p)\Delta p.$$

From (3.4), for any $z \in \mathcal{N}(\tilde{z})$,

$$\Delta\mu + \mu = \rho(z)\bar{\mu}, \qquad \Delta\varepsilon + \varepsilon = \rho(z)\bar{\varepsilon}.$$

Then for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$,

$$\mu + \alpha\Delta\mu = (1 - \alpha)\mu + \alpha\rho(z)\bar{\mu} \in R_{++}, \tag{3.5}$$

$$\varepsilon + \alpha\Delta\varepsilon = (1 - \alpha)\varepsilon + \alpha\rho(z)\bar{\varepsilon} \in R_{+}. \tag{3.6}$$

It follows from the Mean Value Theorem that

$$g_p(\alpha) = \alpha \int_0^1 (\nabla\phi(p + \tau\alpha\Delta p) - \nabla\phi(p))\Delta p \mathrm{d}\tau.$$

Since $\nabla\phi(\cdot)$ is uniformly continuous on $\mathcal{N}(\tilde{z})$ and $\Delta z \to \Delta\tilde{z}$ as $z \to \tilde{z}$, for all $z \in \mathcal{N}(\tilde{z})$,

$$\lim_{\alpha \downarrow 0} \|g_p(\alpha)\|/\alpha = 0.$$

Similarly, note $q = (\varepsilon, x, y)$, $\Delta q = (\Delta\varepsilon, \Delta x, \Delta y)$, $\mathcal{F}(q) = F(x) + \varepsilon x - y$, define $h_q(\alpha) = \mathcal{F}(q + \alpha\Delta q) - \mathcal{F}(q) - \alpha\nabla(\mathcal{F}(q))\Delta q$. By using the Mean Value Theorem, we obtain

$$h_q(\alpha) = \alpha \int_0^1 (\nabla\mathcal{F}(q + \tau\alpha\Delta q) - \nabla\mathcal{F}(q))\Delta q \mathrm{d}\tau.$$

Following the definition of Gateaux differential we have

$$h_q(\alpha) = o(\alpha).$$

Then, from (3.5), (3.6) and the fact that $\rho(z) \le \gamma\Psi(z)^{\frac{1}{2}}$, for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$, we have

$$\begin{aligned}(\mu + \alpha\Delta\mu)^2 &= ((1 - \alpha)\mu + \alpha\rho(z)\bar{\mu})^2 \\ &= (1 - \alpha)^2\mu^2 + \alpha^2\rho^2(z)\bar{\mu}^2 + 2(1 - \alpha)\alpha\rho(z)\mu\bar{\mu} \\ &\le (1 - \alpha)^2\mu^2 + 2\alpha\rho(z)\mu\bar{\mu} + O(\alpha^2) \\ &\le (1 - \alpha)^2\mu^2 + 2\alpha\gamma\Psi(z)^{\frac{1}{2}}\|H(z)\|\bar{\mu} + O(\alpha^2) \\ &= (1 - \alpha)^2\mu^2 + 2\sqrt{2}\alpha\gamma\bar{\mu}\Psi(z) + O(\alpha^2),\end{aligned} \tag{3.7}$$

$$\begin{aligned}(\varepsilon + \alpha\Delta\varepsilon)^2 &= ((1 - \alpha)\varepsilon + \alpha\rho(z)\bar{\varepsilon})^2 \\ &= (1 - \alpha)^2\varepsilon^2 + \alpha^2\rho^2(z)\bar{\varepsilon}^2 + 2(1 - \alpha)\alpha\rho(z)\varepsilon\bar{\varepsilon} \\ &\le (1 - \alpha)^2\varepsilon^2 + 2\alpha\rho(z)\varepsilon\bar{\varepsilon} + O(\alpha^2) \\ &\le (1 - \alpha)^2\varepsilon^2 + 2\alpha\gamma\Psi(z)^{\frac{1}{2}}\|H(z)\|\bar{\varepsilon} + O(\alpha^2) \\ &= (1 - \alpha)^2\varepsilon^2 + 2\sqrt{2}\alpha\gamma\bar{\varepsilon}\Psi(z) + O(\alpha^2),\end{aligned} \tag{3.8}$$

$$\begin{aligned}\|\phi(p + \alpha\Delta p)\|^2 &= \|\phi(p) + \alpha\nabla\phi(p)\Delta p + g_p(\alpha)\|^2 \\ &= \|(1 - \alpha)\phi(p) + g_p(\alpha)\|^2 \\ &= (1 - \alpha)^2\|\phi(p)\|^2 + 2(1 - \alpha)\phi(p)^T g_p(\alpha) + \|g_p(\alpha)\|^2 \\ &= (1 - \alpha)^2\|\phi(p)\|^2 + o(\alpha),\end{aligned} \tag{3.9}$$

$$\begin{aligned}\|\mathcal{F}(q+\alpha\Delta q)\|^2 &= \|\mathcal{F}(q)+\alpha\nabla(\mathcal{F}(q))\Delta q+h_q(\alpha)\|^2\\ &= \|(1-\alpha)\mathcal{F}(q)+h_q(\alpha)\|^2\\ &= (1-\alpha)^2\|\mathcal{F}(q)\|^2+2(1-\alpha)\mathcal{F}(q)^T h_q(\alpha)+\|h_q(\alpha)\|^2\\ &= (1-\alpha)^2\|\mathcal{F}(q)\|^2+o(\alpha).\end{aligned} \tag{3.10}$$

It then follows from (3.7)–(3.10), that for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$. We have

$$\begin{aligned}\Psi(z+\alpha\Delta z) &= \frac{1}{2}\|H(z+\alpha\Delta z)\|^2\\ &= \frac{1}{2}[(\mu+\alpha\Delta\mu)^2+(\varepsilon+\alpha\Delta\varepsilon)^2]+\frac{1}{2}\|\phi(p+\alpha\Delta p)\|^2+\frac{1}{2}\|\mathcal{F}(q+\alpha\Delta q)\|^2\\ &= \frac{1}{2}(1-\alpha)^2(\mu^2+\varepsilon^2+\|\phi(p)\|^2+\|\mathcal{F}(q)\|^2)+\sqrt{2}\alpha\gamma(\bar{\mu}+\bar{\varepsilon})\Psi(z)+o(\alpha)\\ &= (1-\alpha)^2\Psi(z)+\sqrt{2}\alpha\gamma(\bar{\mu}+\bar{\varepsilon})\Psi(z)+o(\alpha)\\ &= [1-2\sigma(1-\gamma)(\bar{\mu}+\bar{\varepsilon})\alpha]\Psi(z)+o(\alpha).\end{aligned} \tag{3.11}$$

Then from equality (3.11) we can find a positive number $\tilde{\alpha} \in (0, 1]$ such that for all $\alpha \in [0, \tilde{\alpha}]$ and all $z \in \mathcal{N}(\tilde{z})$, (3.3) holds. From the above proof we can get Algorithm 3.1 which is well defined at the kth iteration. □

From the above proof we can get the following result directly.

Proposition 3.2. *For each fixed $k \geq 0$ if $z_k \in \Omega$, then for any $\alpha \in [0, 1]$ such that*

$$\Psi(z_k+\alpha\Delta z_k) \leq [1-2\sigma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k), \tag{3.12}$$

we have $z_k + \alpha\Delta z_k \in \Omega$.

Proof. We prove this proposition by considering the following two cases:

(i) If $\Psi(z_k) > 1$. Then $\rho_k = \gamma$. It therefore follows from $z_k \in \Omega$ and $\rho(z) = \gamma \min\{1, \Psi(z)\} \leq \gamma$ for any $z \in R_{++} \times R_+ \times R^n \times R^n$ that for all $\alpha \in [0, 1]$, we have

$$\begin{aligned}\mu_k+\alpha\Delta\mu_k-\rho(z_k+\alpha\Delta z_k)\bar{\mu} &\geq (1-\alpha)\mu_k+\alpha\rho_k\bar{\mu}-\gamma\bar{\mu}\\ &\geq (1-\alpha)\rho_k\bar{\mu}+\alpha\rho_k\bar{\mu}-\gamma\bar{\mu}\\ &= \rho_k\bar{\mu}-\gamma\bar{\mu}=0,\end{aligned} \tag{3.13}$$

$$\begin{aligned}\varepsilon_k+\alpha\Delta\varepsilon_k-\rho(z_k+\alpha\Delta z_k)\bar{\varepsilon} &\geq (1-\alpha)\varepsilon_k+\alpha\rho_k\bar{\varepsilon}-\gamma\bar{\varepsilon}\\ &\geq (1-\alpha)\rho_k\bar{\varepsilon}+\alpha\rho_k\bar{\varepsilon}-\gamma\bar{\varepsilon}\\ &= \rho_k\bar{\varepsilon}-\gamma\bar{\varepsilon}=0.\end{aligned} \tag{3.14}$$

Therefore, in this case $z_k + \alpha\Delta z_k \in \Omega$.

(ii) If $\Psi(z_k) \leq 1$. Then for any $\alpha \in [0, 1]$ satisfying (3.12), we have

$$\Psi(z_k+\alpha\Delta z_k) \leq [1-2\sigma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k) \leq 1. \tag{3.15}$$

So, for any $\alpha \in [0, 1]$, satisfying (3.12), we have

$$\rho(z_k+\alpha\Delta z_k) = \gamma\Psi(z_k+\alpha\Delta z_k).$$

Hence, again because $z_k \in \Omega$, by using the first inequality in (3.15), for any $\alpha \in [0, 1]$ satisfying (3.12) we have

$$\begin{aligned}\mu_k+\alpha\Delta\mu_k-\rho(z_k+\alpha\Delta z_k)\bar{\mu} &= (1-\alpha)\mu_k+\alpha\rho_k\bar{\mu}-\gamma\Psi(z_k+\alpha\Delta z_k)\bar{\mu}\\ &\geq (1-\alpha)\rho_k\bar{\mu}+\alpha\rho_k\bar{\mu}-\gamma[1-2\sigma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k)\bar{\mu}\\ &= \rho_k\bar{\mu}-\gamma[1-2\sigma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k)\bar{\mu}\\ &= \gamma\Psi(z_k)\bar{\mu}-\gamma[1-2\sigma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k)\bar{\mu}\\ &= [2\sigma\gamma(1-\gamma(\bar{\mu}+\bar{\varepsilon}))\alpha]\Psi(z_k)\bar{\mu} \geq 0.\end{aligned} \tag{3.16}$$

Similar arguments hold for $\varepsilon_k + \alpha\Delta\varepsilon_k$,

$$\varepsilon_k+\alpha\Delta\varepsilon_k-\rho(z_k+\alpha\Delta z_k)\bar{\varepsilon} = (1-\alpha)\varepsilon_k+\alpha\rho_k\bar{\varepsilon}-\gamma\Psi(z_k+\alpha\Delta z_k)\bar{\varepsilon} \geq 0. \tag{3.17}$$

Therefore, in this case $z_k + \alpha\Delta z_k \in \Omega$. Thus, by combining the above two cases, we have proved that for all $\alpha \in [0, 1]$ satisfying (3.12) $z_k + \alpha\Delta z_k \in \Omega$. This completes our proof. □

By combining Proposition 3.1 and Proposition 3.2, we have

Proposition 3.3. *Let $\{z^k = (\mu_k, \varepsilon_k, x_k, y_k)\}$ be the iteration sequence generated by Algorithm 3.1, if for each fixed $k \geq 0$, $z^k \in \Omega$, then $z_{k+1} \in \Omega$.*

Proposition 3.4. *Since $\nabla H(z_k)$ is invertible for every $k \geq 0$, with $z^k \in \Omega$, then for an infinite sequence $\{z_k\}$ generated by Algorithm 3.1, we have $\{z_k\} \in \Omega$.*

Proof. First, since $z_0 = (\mu_0, \varepsilon_0, x_0, y_0) \in \Omega$, we have from Proposition 3.3 that z_1 is well defined and $z_1 \in \Omega$. Then, by repeatedly resorting to Proposition 3.3 we can prove that an infinite sequence $\{z_k\}$ is generated. □

4. Convergence analysis

In this section, we consider the convergence and convergent rate of the proposed algorithm. First, we introduce the following assumption.

Assumption 4.1. The solution set $S = \{x \in \mathcal{K}, F(x) \in \mathcal{K}, \langle x, F(x)\rangle = 0\}$ of SOCCP is nonempty and bounded.

Similar to the proof of Lemma 5.3 in [10], we have the following lemma.

Lemma 4.1. *Suppose that F is a continuously monotone function, and H is defined by (2.5). Let $\{z^k = (\mu_k, \varepsilon_k, x^k, y^k)\}$ be the iteration sequence generated by Algorithm 3.1. Then level set*

$$L(z^0) = \{z \mid \|H(z)\| \leq \|H(z^0)\|\} \tag{4.1}$$

is bounded and $\{z^k\} \subset L(z^0)$.

Theorem 4.1. *Assume that Assumption 4.1 holds and $\{z_k\}$ is an infinite sequence generated by the algorithm. Then the sequence $\{z^k\}$ is bounded and each accumulation point z^* of $\{z_k\}$ is a solution of Eq. (2.5), i.e., $H(z^*) = 0$.*

Proof. Obviously, by Lemma 4.1, $\{z^k\}$ is bounded. Furthermore, it follows from Proposition 3.4 that the infinite sequence $\{z_k\}$ generated by Algorithm 3.1 is in the set Ω. From the design of Algorithm 3.1, we have $\Psi(z_{k+1}) \leq \Psi(z_k)$ for all $k \geq 0$. Hence both the sequences $\{\Psi(z_k)\}$ and $\{\rho(z_k)\}$ are monotonically decreasing. Since $\Psi(z_k), \rho(z_k) \geq 0$ $(k \geq 0)$, there exist $\Psi^*, \rho^* \geq 0$ such that $\Psi(z_k) \to \Psi^*$ and $\rho(z_k) \to \rho^*$ as $k \to \infty$.

If $\Psi^* = 0$ and $\{z_k\}$ has an accumulation point z^*, then from the continuity of $\Psi(\cdot)$ and $\rho(\cdot)$ we obtain $\Psi(z^*) = 0$ and $\rho(z^*) = 0$. Then we obtain the desired result.

Suppose that $\Psi^* > 0$, $z^* = (\mu^*, \varepsilon^*, x^*, y^*)$ is an accumulation point of $\{z_k\}$. We may assume that $\{z_k\}$ converges to z^*. It is easy to see that $\Psi^* = \Psi(z^*)$, $\rho^* = \rho(z^*)$.

From Proposition 3.1 there exist a closed neighborhood $\mathcal{N}(z^*)$ of z^* and a positive number $\tilde{\alpha} \in (0, 1]$ such that for any $z \in \mathcal{N}(z^*)$ and all $\alpha \in [0, \tilde{\alpha}]$, $\nabla H(z)$ is invertible and (3.3) holds. Therefore, for a nonnegative integer l such that $\delta^l \in (0, \tilde{\alpha}]$, we have

$$\Psi(z_k + \delta^l \Delta z_k) \leq [1 - 2\sigma(1 - \gamma(\bar{\mu} + \bar{\varepsilon}))\delta^l]\Psi(z_k),$$

for all sufficiently large k. Then, for every sufficiently large k, we see that $l^k \leq l$ and hence $\delta^{l_k} \geq \delta^l$. Then,

$$\Psi(z_{k+1}) \leq [1 - 2\sigma(1 - \gamma(\bar{\mu} + \bar{\varepsilon}))\delta^{l_k}]\Psi(z_k) \leq [1 - 2\sigma(1 - \gamma(\bar{\mu} + \bar{\varepsilon}))\delta^l]\Psi(z_k)$$

for all sufficiently large k. This contradicts the fact that the sequence $\{\Psi(z_k)\}$ converges to $\Psi^* > 0$. So, we complete our proof. □

Theorem 4.1 discusses the global convergence of the algorithm. Now, we will analyze the rate of convergence of Algorithm 3.1. Similar to the proof of Theorem 3.2 in [11], we have the following result.

Lemma 4.2. *Let $H(z) := H(\mu, \varepsilon, x, y)$ be defined by (2.5). Then $H(z)$ is strongly semismooth at any $z := (\mu, \varepsilon, x, y) \in R \times R \times R^n \times R^n$. Therefore, we have*

$$\|H(z + h) - H(z) - Vh\| = O(\|h\|^2), \quad \forall V \in \partial H(z + h),\ h \to 0. \tag{4.2}$$

Theorem 4.2. *Suppose that $z^* = (\mu^*, \varepsilon^*, x^*, y^*)$ is an accumulation point of the iteration sequence $\{z_k\}$ generated by Algorithm 3.1. If all $V \in \partial H(z^*)$ are nonsingular at z^*, where ∂H stands for the generalized Jacobian of H in the sense of Clarke. Then $\{z_k\}$ converges to z^* quadratically, i.e., $\|z_{k+1} - z^*\| = O(\|z_k - z^*\|^2)$.*

Proof. First, from Theorem 4.1 that z^* is a solution of $H(z) = 0$. Then from Proposition 3.1 of [12], for all z_k sufficiently close to z^*, we obtain

$$\|\nabla H(z_k)^{-1}\| = O(1).$$

Notice that H is strong semismooth at z^*, from (3.2) for z_k sufficiently close to z^*, we have

$$\begin{aligned}\|z_k + \Delta z_k - z^*\| &= \|z_k + \nabla H(z_k)^{-1}(-H(z_k) + \rho_k \bar{z}) - z^*\| \\ &= \|\nabla H(z_k)^{-1}[\nabla H(z_k)(z_k - z^*) + (-H(z_k) + \rho_k \bar{z})]\| \\ &\leq \|\nabla H(z_k)^{-1}\| \, \|\nabla H(z_k)(z_k - z^*) + (-H(z_k) + \rho_k \bar{z})\| \\ &= O(\|H(z_k) - H(z^*) - \nabla H(z_k)(z_k - z^*)\| + \rho_k \bar{z}) \\ &= O(\|z_k - z^*\|^2) + O(\Psi(z_k)). \end{aligned} \tag{4.3}$$

Since H is strongly semismooth at z^*, H is locally Lipschitz continuous near z^*, for all z_k close to z^*,

$$\Psi(z_k) = \frac{1}{2}\|H(z_k)\|^2 = O(\|z_k - z^*\|^2). \tag{4.4}$$

Therefore, from (4.3) and (4.4), for all z_k sufficiently close to z^*,

$$\|z_k + \Delta z_k - z^*\| = O(\|z_k - z^*\|^2). \tag{4.5}$$

By following the proof of Theorem 3.1 of [13], for all z_k sufficiently close to z^*, we have

$$\|z_k - z^*\| = O(\|H(z_k) - H(z^*)\|) \tag{4.6}$$

Hence, for all z_k sufficiently close to z^*, we have

$$\begin{aligned}\Psi(z_k + \Delta z_k) &= \frac{1}{2}\|H(z_k + \Delta z_k)\|^2 = O(\|z_k + \Delta z_k - z^*\|^2) \\ &= O(\|z_k - z^*\|^4) = O(\|H(z_k) - H(z^*)\|^4) \\ &= O(\Psi(z_k)^2). \end{aligned} \tag{4.7}$$

Therefore, for all z_k sufficiently close to z^* we have $z_{k+1} = z_k + \Delta z_k$, which, together with (4.5), implies that $\|z_{k+1} - z^*\| = O(\|z_k - z^*\|^2)$. $\square$

5. Numerical experiments

In this section, we consider some numerical examples to evaluate the efficiency of Algorithm 3.1. In our experiment, we choose parameters $\mu_0 = 0.5$, $\varepsilon_0 = 0.5$, $\delta = 0.8$, $\sigma = 0.25$ and $\gamma = 0.5$. We employed $\Psi(z^k) < 10^{-20}$ as the termination criterion. All the program codes were written in MATLAB and run in MATLAB 7.6 environment. All numerical experiments were done at a PC with Celeron(R) D CPU of 3.06 GHz and RAM of 512 MB.

In the tables of test results, DIM denotes the dimension of the problem (the dimension of the variable x), SP denotes the starting point of (x^0, y^0), Iter denotes the iterative number, FV denotes the final value of the merit function $\Psi(z^k)$ when the algorithm terminates, and CPU records the CPU time in second for solving each problem. In the following, we give a detailed description of the tested problems.

Example 5.1. The second-order cone linear complementarity problem on $\mathcal{K}^3$. The test function F and the second-order cone $\mathcal{K}$ are given as follows:

$$F(x) = \begin{pmatrix} 21 & -9 & 18 \\ -9 & 4 & -7 \\ 18 & -7 & 19 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \\ 1 \end{pmatrix}, \qquad \mathcal{K} := \mathcal{K}^3.$$

In this example, $F : R^3 \to R^3$ is an affine function whose Jacobian is symmetric and positive semidefinite but not positive definite. This example has one solution $x^* \approx (0.183606, -0.154346, -0.099440)^T$. We test this problem by the different starting points and the test results are listed in Table 1.

Example 5.2. The second-order cone linear complementarity problem on $\mathcal{K}^5$. The test function F and the second-order cone $\mathcal{K}$ are given as follows:

$$F(x) = \begin{pmatrix} 15x_1 - 5x_2 - x_3 + 4x_4 - 5x_5 \\ 5x_2 + x_5 \\ -x_1 - 3x_2 + 8x_3 + 2x_4 - 3x_5 \\ 2x_1 - 4x_2 + 2x_3 + 9x_4 - 4x_5 \\ -5x_2 + 10x_5 - 1 \end{pmatrix}, \qquad \mathcal{K} := \mathcal{K}^5.$$

This example has one solution $x^* \approx (0.049185, -0.0030997, 0.0096024, 0.0031883, 0.048033)^T$. We test this problem by the different starting points and the test results are listed in Table 2.

Table 1
Numerical results for Example 5.1.

$\text{SP}((x^0, y^0)^\top)$	IN	FV	CPU
$(0, 0, 0; 0, 0, 0)$	8	1.5544×10^{-27}	0.0185
$(1, 1, 1; 1, 1, 1)$	8	4.5790×10^{-27}	0.0187
$(-1, -1, -1; -1, -1, -1)$	8	9.3277×10^{-27}	0.0185
$(10, 10, 10; 10, 10, 10)$	8	2.9065×10^{-26}	0.0186
$(-10, -10, -10; -10, -10, -10)$	8	4.3718×10^{-27}	0.0186
$(10^3, 10^3, 10^3; 10^3, 10^3, 10^3)$	10	2.8202×10^{-29}	0.0227
$(10^6, 10^6, 10^6; 10^6, 10^6, 10^6)$	10	1.4288×10^{-28}	0.0228
$(10^9, 10^9, 10^9; 10^9, 10^9, 10^9)$	10	1.5777×10^{-30}	0.0236

Table 2
Numerical results for Example 5.2.

$\text{SP}((x^0, y^0)^\top)$	IN	FV	CPU
$(0, \dots, 0; 0, \dots, 0)$	7	1.5054×10^{-25}	0.0168
$(1, \dots, 1; 1, \dots, 1)$	7	1.2075×10^{-21}	0.0171
$(-1, \dots, -1; -1, \dots, -1)$	7	7.1501×10^{-25}	0.0167
$(10, \dots, 10; 10, \dots, 10)$	8	6.7817×10^{-26}	0.0192
$(-10, \dots, -10; -10, \dots, -10)$	8	8.3138×10^{-23}	0.0191
$(10^3, \dots, 10^3; 10^3, \dots, 10^3)$	8	1.4644×10^{-21}	0.0195
$(10^6, \dots, 10^6; 10^6, \dots, 10^6)$	8	1.7186×10^{-21}	0.0196
$(10^9, \dots, 10^9; 10^9, \dots, 10^9)$	8	1.7190×10^{-21}	0.0201

Table 3
Numerical results for Example 5.3.

$\text{SP}((x^0, y^0)^\top)$	IN	FV	CPU
$(0, 0, 0; 0, 0, 0)$	7	3.7855×10^{-27}	0.0171
$(1, 1, 1; 1, 1, 1)$	7	2.6096×10^{-26}	0.0166
$(-1, -1, -1; -1, -1, -1)$	7	4.4168×10^{-26}	0.0169
$(10, \dots, 10; 10, \dots, 10)$	8	2.9511×10^{-41}	0.0190
$(-10, -10, -10; -10, -10, -10)$	10	4.9304×10^{-31}	0.0234
$(30, 30, 30; 30, 30, 30)$	10	1.6197×10^{-25}	0.0233
$(10^2, 10^2, 10^2; 10^2, 10^2, 10^2)$	13	1.0570×10^{-25}	0.0308
$(10^3, 10^3, 10^3; 10^3, 10^3, 10^3)$	19	3.4301×10^{-28}	0.0431
$(-10^3, -10^3, -10^3; -10^3, -10^3, -10^3)$	18	5.3288×10^{-30}	0.0426

Example 5.3. The second-order cone nonlinear complementarity problem on $\mathcal{K}^3$. The test function F and the second-order cone $\mathcal{K}$ are given as follows:

$$F(x) = \begin{pmatrix} 0.07x_1^3 - 4 \\ 0.04x_2^3 - 3.93 \\ 0.03x_3^3 - 5.72 \end{pmatrix}, \qquad \mathcal{K} := \mathcal{K}^3.$$

In this example, $F : R^3 \to R^3$ is a strictly monotone function comprised of cubic and constant terms. It is easy to see that its Jacobian $F'(x) = \text{diag}(0.21x_1^2, 0.12x_2^2, 0.09x_3^2)$ is only positive semidefinite, that is, F is only monotone but not strongly monotone. This example has one solution $x^* = (5, 3, 4)^T$. We test this problem by the different starting points and the test results are listed in Table 3.

Example 5.4. The second-order cone nonlinear complementarity problem on $\mathcal{K}^3 \times \mathcal{K}^2$. The test function F and the second-order cone K are given as follows:

$$F(x) = \begin{pmatrix} 24(2x_1 - x_2)^3 + e^{x_1 - x_3} - 4x_4 + x_5 \\ -12(2x_1 - x_2)^3 + 3(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 6x_4 - 7x_5 \\ -e^{x_1 - x_3} + 5(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 3x_4 + 5x_5 \\ 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \end{pmatrix}, \qquad \mathcal{K} := \mathcal{K}^3 \times \mathcal{K}^2.$$

In this example, F is the function which appears in the KKT conditions for the second-order cone programming (SOCP):

$$\min e^{z_1 - z_3} + 3(2z_1 - z_2)^4 + \sqrt{1 + (3z_2 + 5z_3)^2}$$
$$\text{s.t.} \begin{pmatrix} z_1 \\ z_2 \\ z_2 \end{pmatrix} \in \mathcal{K}^3, \quad \begin{pmatrix} 4 & 6 & 3 \\ -1 & 7 & -5 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix} \in \mathcal{K}^2.$$

Table 4
Numerical results for Example 5.4.

$\mathrm{SP}((x^0, y^0)^\top)$	IN	FV	CPU
$(0, \dots, 0; 0, \dots, 0)$	11	1.6662×10^{-25}	0.0566
$(1, \dots, 1; 0, \dots, 0)$	12	6.5790×10^{-31}	0.0577
$(0, \dots, 0; 1, \dots, 1)$	10	1.5708×10^{-24}	0.0506
$(1, \dots, 1; 1, \dots, 1)$	11	9.5504×10^{-22}	0.0492
$(-1, \dots, -1; -1, \dots, -1)$	14	8.4886×10^{-25}	0.0751
$(10, \dots, 10; 10, \dots, 10)$	16	3.0860×10^{-30}	0.0739
$(-10, \dots, -10; -10, \dots, -10)$	17	8.8698×10^{-24}	0.0859
$(10^2, \dots, 10^2; 10^2, \dots, 10^2)$	25	8.1396×10^{-28}	0.1146
$(-10^2, \dots, -10^2; -10^2, \dots, -10^2)$	18	4.5093×10^{-23}	0.0896

From the convexity of the objective function, we can easily prove that F is monotone. This example has one solution $x^* \approx (0.232402, -0.073079, 0.220614, 0.533903, -0.533903)^T$. We test this problem by the different starting points and the test results are listed in Table 4.

References

[1] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM J. Optim. 12 (2002) 436–460.
[2] S. Hayashi, N. Yamashita, M. Fukushima, A combined smoothing and regularization method for monotone second-order cone complementarity problems, SIAM J. Optim. 15 (2005) 593–615.
[3] J.-S. Chen, Two class of merit functions for the second-order cone complementarity problem, Math. Methods Oper. Res. 64 (2006) 495–519.
[4] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Math. Program. 104 (2005) 293–327.
[5] S. Pan, J.-S. Chen, A damped Gauss–Newton method for the second-order cone complementarity problem, Appl. Math. Optim. 59 (2009) 293–318.
[6] X.D. Chen, D. Sun, J. Sun, Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems, Comput. Optim. Appl. 25 (2003) 39–56. 2003.
[7] C.-K. Sim, G. Zhao, A note on treating a second order cone program as a special case of a semidefinite program, Math. Program. 102 (2005) 609–613.
[8] U. Faraut, A. Korányi, Analysis on Symmetric Cones, in: Oxford Mathematical Monographs, Oxford University Press, New York, 1994.
[9] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, SIAM J. Matrix Anal. Appl. 17 (1996) 851–868.
[10] Z.-H. Huang, T. Ni, Smoothing algorithms for complementarity problems over symmetric cones, Comput. Optim. Appl. 45 (2010) 557–579.
[11] D. Sun, J. Sun, Strong semismoothness of the Fischer–Burmeister SDC and SOC complementarity functions, Math. Program. A 103 (2005) 575–581.
[12] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Program. 58 (1993) 353–367.
[13] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993) 227–244.